3K

УДК 517.958:531.33

Ю.В. Крак, І.О. Стеля
Київський національний університет імені Тараса Шевченка, м. Київ, Україна
krak@unicyb.kiev.ua, igor.stelia@gmail.com

Моделювання мовного апарату людини в задачах комп'ютерного синтезу

У статті отримала розвиток методика розв'язання проблеми синтезу мовної інформації, що дозволило розв'язати задачу штучного утворення звуків голосу людини на основі спільного використання фізичних моделей голосового джерела та мовного тракту. На базі розроблених алгоритмів створено програмно-алгоритмічний комплекс для дослідження впливу параметрів моделей у системі «голосове джерело – мовний тракт» на мовотворення. Для акустичного рівняння Клейна – Гордона розв'язана обернена задача відновлення параметрів мовного тракту за вимірюваним сигналом на виході.

Вступ

Розвиток комп'ютерних мовних синтезаторів привів до розробки декількох підходів до їх побудови, а саме:

– синтез на основі конкатенації відрізків записаних мовленнєвих сигналів, що зберігаються у мовленнєвих базах даних;

– формантний синтез, що не використовує зразки людської мови для синтезу. Замість цього синтезована мова створюється з використанням акустичних моделей. Параметри, такі, як набір значень частот основних формант та рівні шуму, змінюються в часі та створюють сигнал штучної мови;

– артикуляторний синтез, що створює штучну мову шляхом безпосереднього моделювання людського мовного апарату. Цей метод являє собою синтез мови на базі обчислювальних методів, заснованих на моделях людського мовного апарату і артикуляційних процесах, що відбуваються у ньому. Моделі мовного апарату, як правило, включають модель мовного тракту і модель голосових зв'язок.

Для побудови та дослідження математичних моделей голосового джерела, що моделюють коливні процеси, та моделей мовного тракту, що моделюють розповсюдження акустичних хвиль, використовуються чисельні методи. Для задачі відновлення параметрів мовного тракту за вимірюваним сигналом на виході використовується математичний апарат розв'язання обернених задач. Проблемі розв'язання мовних обернених задач присвячені, зокрема, роботи [1], [2].

Метою дослідження є комп'ютерне відтворення звуків, що утворює людина, на основі спільного використання фізичних моделей голосового джерела та мовного тракту.

Модель голосових зв'язок Ішізаки – Фланагана

У моделі Ішізаки і Фланагана [3] кожна з голосових зв'язок описується двома масами, що зв'язані пружинами як зі стінками, так і між собою. Більш того, припускається, що зв'язки є двосторонньо симетричними. Маси, що моделюють зв'язки, здійснюють коливання у поперечному до руху повітря напрямку. Система рівнянь

для двох мас, що коливаються, записується у вигляді:

$$m_1 \frac{d^2 x_1(t)}{dt^2} + r_1 \frac{dx_1(t)}{dt} + k_1(x_1(t) - x_{01}) + k_c(x_1(t) - x_2(t)) = l_g d_1 p_{m1}(t),$$
$$m_2 \frac{d^2 x_2(t)}{dt^2} + r_2 \frac{dx_2(t)}{dt} + k_2(x_2(t) - x_{02}) - k_c(x_1(t) - x_2(t)) = l_g d_2 p_{m2}(t), \quad (1)$$

де m_1 і m_2 – маси; $x_1(t)$, $x_2(t)$ – зміщення мас m_1 і m_2; x_{01}, x_{02} – початкове положення мас m_1 і m_2; t – час; r_1 і r_2 – коефіцієнти демпфування; k_1, k_2 – пружність пружин для мас m_1 і m_2; k_c – пружність пружини, що з'єднує маси m_1 і m_2; d_1, d_2 – товщина мас m_1 і m_2; l_g – діюча довжина голосових зв'язок; $l_g d_1$, $l_g d_2$ – поверхні мас m_1 і m_2, на які діють тиски $p_{m1}(t)$ і $p_{m2}(t)$ відповідно.

Розподіл тиску в голосовій щілині апроксимується послідовними дискретними кроками p_{ij} на кожному j-му кінці кожної i-ї маси. Перше падіння тиску $p_s - p_{11}(t)$ виводиться з рівняння Бернуллі для ідеальної рідини у стаціонарному режимі. Вздовж кожної з мас падіння тиску $p_{11}(t) - p_{12}(t)$ та $p_{21}(t) - p_{22}(t)$ визначається в'язкими втратами і пропорційне зсувовій в'язкості повітря. Падіння тиску на стику мас $p_{12}(t) - p_{21}(t)$ еквівалентне зміні кінетичної енергії в одиниці об'єму газу. І остаточно різке розширення на верхньому кінці голосової щілини призводить до відновлення тиску до атмосферного тиску p. Таким чином, система рівнянь для змін тиску записується у вигляді [3]:

$$p_s - p_{11}(t) = 0.69\rho \frac{u_g^2(t)}{A_{g1}^2(t)} + \int_0^{l_c} \frac{\rho}{A_c(x)} dx \cdot \frac{du_g}{dt},$$
$$p_{11}(t) - p_{12}(t) = 12\nu d_1 \frac{l_g^2 u_g(t)}{A_{g1}^3(t)} + \frac{\rho d_1}{A_{g1}} \cdot \frac{du_g}{dt},$$
$$p_{12}(t) - p_{21}(t) = \frac{1}{2}\rho u_g^2(t)\left(\frac{1}{A_{g2}^2(t)} - \frac{1}{A_{g1}^2(t)}\right), \quad (2)$$
$$p_{21}(t) - p_{22}(t) = 12\nu d_2 \frac{l_g^2 u_g(t)}{A_{g2}^3(t)} + \frac{\rho d_2}{A_{g2}} \cdot \frac{du_g}{dt},$$
$$p_{22}(t) - p = \frac{1}{2}\rho \frac{u_g^2(t)}{A_{g2}^2(t)} \left[2\frac{A_{g2}(t)}{A_1}\left(1 - \frac{A_{g2}(t)}{A_1}\right) \right],$$

де ρ – густина повітря; ν – зсувова в'язкість повітря; A_1 – площа голосового тракту на вході; A_{gi} – площа голосової щілини під i-ю масою; $u_g(t)$ – потік повітря;

$$A_{gi}(t) = (A_{g0i} + 2l_g x_i(t)),\ i = 1, 2,$$

де $x_1(t) \geq x_{01}$, $x_2(t) \geq x_{02}$; A_{g01}, A_{g02} – залишкові площі в момент змикання голосових зв'язок.

Для визначення значень тиску $p_{m1}(t)$, $p_{m2}(t)$ використовуються співвідношення:

$$p_{m1}(t) = \frac{1}{2}(p_{11}(t) + p_{12}(t)),\ p_{m2}(t) = \frac{1}{2}(p_{21}(t) + p_{22}(t)). \quad (3)$$

Шуканим розв'язком системи (1) – (3) є функція $u_g(t)$, яка визначає потік повітря на виході з голосової щілини.

Для побудови чисельного алгоритму розв’язання систем (1) та (2) використаємо позначення для зміщень та похідних:

$$V_1(t) = x_1(t)\,,\ V_2(t) = \frac{dx_1}{dt}\,,\ V_3(t) = x_2(t)\,,\ V_4(t) = \frac{dx_2}{dt}$$

і перепишемо систему (1) у вигляді:

$$\begin{gathered}\frac{dV_1(t)}{dt} = \varphi_1(V_2(t))\,,\\ \frac{dV_2(t)}{dt} = \varphi_2(V_1(t), V_2(t), V_3(t))\,,\\ \frac{dV_3(t)}{dt} = \varphi_3(V_4(t))\,,\\ \frac{dV_4(t)}{dt} = \varphi_4(V_1(t), V_3(t), V_4(t))\,,\ t \in (0, T]\,.\end{gathered} \tag{4}$$

До цієї системи додаються початкові умови:

$$V_1(0) = x_{01}\,,\ V_2(0) = x'_{01}\,,\ V_3(0) = x_{02}\,,\ V_4(0) = x'_{02}\,.$$

Для розв’язання системи (4) будемо використовувати метод Кутта – Мерсона [4]. Для використання чисельного методу здійснюється дискретизація відрізка інтегрування системи (1). Будемо використовувати рівномірне розбиття відрізка $0 \le t \le T$: $x_{i+1} = x_i + h$, $i = 0...N_T$, $N_T = T / h$.

Для побудови чисельного алгоритму розв’язання нелінійної системи (3) запишемо її у вигляді:

$$\begin{gathered}f_1(p_{11}, u_g) = 0,\\ f_2(p_{11}, p_{12}, u_g) = 0,\\ f_3(p_{12}, p_{21}, u_g) = 0,\\ f_4(p_{21}, p_{22}, u_g) = 0,\\ f_5(p_{22}, u_g) = 0\,.\end{gathered} \tag{5}$$

Систему (5) будемо розглядати як операторне рівняння в деякому просторі. Нехай

$$\begin{gathered}P = (p_{11}, p_{12}, p_{21}, p_{22}, u_g)^T\,,\\ F(P) = (f_1(P), f_2(P), f_3(P), f_4(P), f_5(P))^T\,,\\ F(P) = 0\,,\end{gathered} \tag{6}$$

де F – нелінійне відображення.

Для розв’язання системи (6) будемо використовувати метод Ньютона [4]. Для цього запишемо розвинення для $f_i(P)$, $i = \overline{1,5}$ в ряд Тейлора. Тоді, згідно з методом Ньютона, наступне наближення до розв’язку нелінійної системи (6) знаходиться як розв’язок лінійної системи:

$$F'(P^s)P^{s+1} = F'(P^s)P^s - F(P^s),\ \text{де}\ s = 1,2,3,...,. \tag{7}$$

Для розв’язання лінійної системи (7) на кожному кроці методу Ньютона будемо використовувати ітераційний метод послідовної верхньої релаксації (ПВР), ітераційна схема якого має вигляд:

$$(D - \omega L)P_{k+1}^{s+1} = [(1-\omega)D + \omega U]P_k^{s+1} + \omega\left(F'(P_k^{s+1})P_k^{s+1} - F(P_k^{s+1})\right),$$

де матриця A подана у вигляді суми діагональної, нижньої трикутної та верхньої трикутної матриці $F'(P^s) = D - L - U$, ω – релаксаційний параметр.

Алгоритм розв'язання системи (1) – (3) будується таким чином: нехай відомий розв'язок у момент часу t_k: $V_1(t_k), V_2(t_k), V_3(t_k), V_4(t_k), p_{11}(t_k),\ p_{21}(t_k),\ p_{21}(t_k),\ p_{22}(t_k)$, $u_g(t_k)$, знаходження розв'язку системи (4) на ($k+1$)-у кроці явного методу Кутта – Мерсона об'єднується в єдину ітераційну процедуру з розв'язанням системи (3) методом послідовної верхньої релаксації [5], згідно з яким перехід на ($s+1$)-у ітерацію здійснюється за формулами:

$$
\begin{aligned}
&V_1^{s+1}(t_{k+1}) = KM(V_2(t_k)),\\
&V_2^{s+1}(t_{k+1}) = KM(V_1(t_k), V_2(t_k), V_3(t_k); p_{11}^s(t_{k+1}), p_{12}^s(t_{k+1})),\\
&V_3^{s+1}(t_{k+1}) = KM(V_4(t_k)),\\
&V_4^{s+1}(t_{k+1}) = KM(V_1(t_k), V_3(t_k), V_4(t_k); p_{21}^s(t_{k+1}), p_{22}^s(t_{k+1})),
\end{aligned}
$$

де $KM()$ позначено процедуру методу Кутта – Мерсона, а тиски знаходяться методом ПВР наступним чином:

$$p_{11}^{s+1}(t_{k+1}) = p_{11}^s + \omega \frac{f_1\left(p_{11}^s, u_g^s, V_1^{s+1}(t_{k+1})\right)}{\partial f_1 / \partial p_{11}},$$

$$p_{12}^{s+1}(t_{k+1}) = p_{12}^s + \omega \frac{f_2\left(p_{11}^s(t_{k+1}), p_{12}^s(t_{k+1}), u_g^s(t_{k+1}), V_1^{s+1}(t_{k+1})\right)}{\partial f_2 / \partial p_{12}},$$

$$p_{21}^{s+1}(t_{k+1}) = p_{21}^s + \omega \frac{f_3\left(p_{12}^{s+1}(t_{k+1}), p_{21}^s(t_{k+1}), u_g^s(t_{k+1}), V_1^{s+1}(t_{k+1}), V_3^{s+1}(t_{k+1})\right)}{\partial f_3 / \partial p_{21}},$$

$$p_{22}^{s+1}(t_{k+1}) = p_{22}^s + \omega \frac{f_4\left(p_{21}^{s+1}(t_{k+1}), p_{22}^s(t_{k+1}), u_g^s(t_{k+1}), V_3^{s+1}(t_{k+1})\right)}{\partial f_4 / \partial p_{22}},$$

$$u_g^{s+1}(t_{k+1}) = u_g^s + \omega \frac{f_5\left(p_{22}^{s+1}(t_{k+1}), u_g^s(t_{k+1}), V_3^{s+1}(t_{k+1})\right)}{\partial f_5 / \partial u_g}.$$

Ітераційна процедура вважається закінченою за виконання умови: $|u_g^{s+1}(t_k) - u_g^s(t_k)| < \varepsilon$, де $\varepsilon > 0$ – мала величина. Після цього здійснюється перехід на наступний крок методу Кутта – Мерсона. Кількість кроків методу Кутта – Мерсона залежить від часу моделювання T та необхідної точності розрахунків.

За розробленими алгоритмами проведена серія чисельних експериментів, результати яких наведені на рис. 1.

Здійснене порівняння потоку повітря, обчисленого за двомасовою моделлю, та потоку повітря, отриманого за аналітичною R-моделлю [6]:

$$
u_g(t) = \begin{cases}
\alpha\left(3\left(\dfrac{t}{t_p}\right)^2 - 2\left(\dfrac{t}{t_p}\right)^3\right), & 0 \le t \le t_p,\\[2ex]
\alpha\left(1 - \left(\dfrac{t - t_p}{t_p}\right)^2\right), & t_p \le t \le t_p + t_n \le t_0,
\end{cases}
$$

де α – амплітуда, t_p – час відкриття, t_n – час закриття, t_0 – період. Результати порівняння моделей наведені на рис. 2.

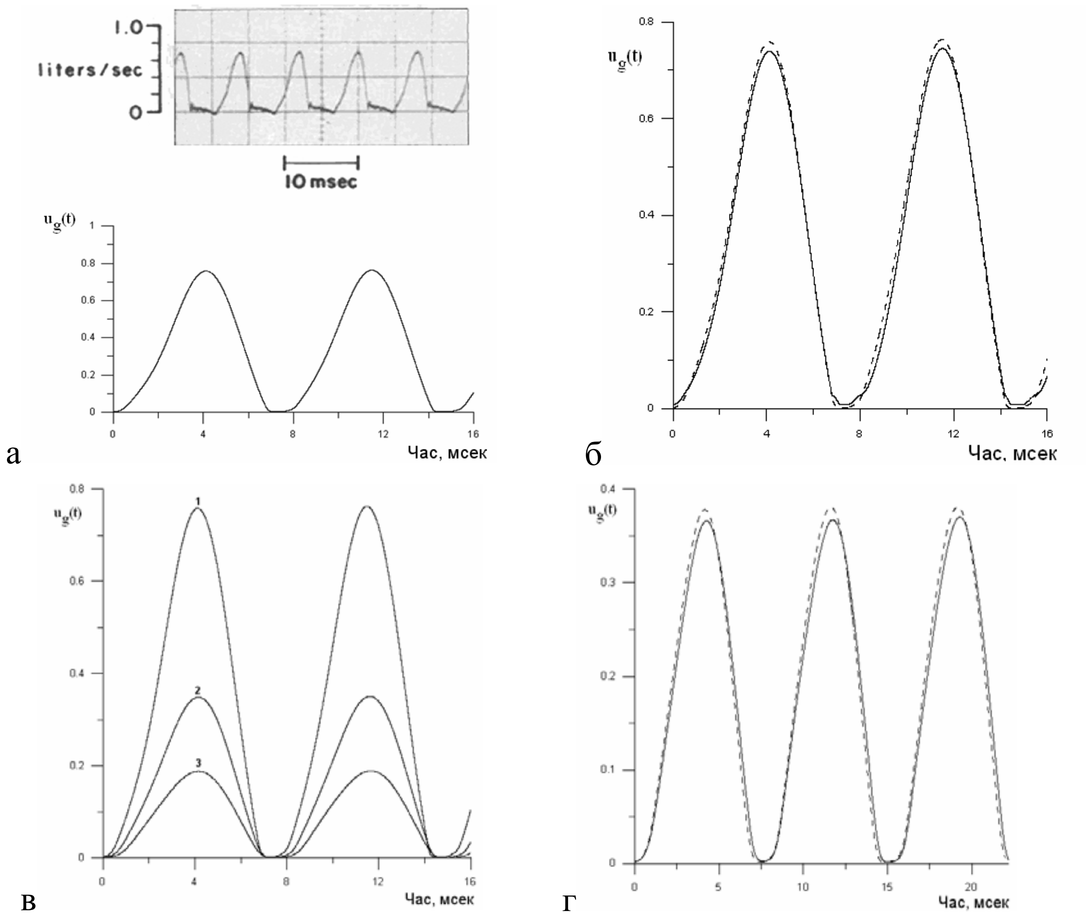


Рисунок 1 – Потік повітря u_g на виході з голосових зв'язок, виміряний за допомогою зворотної фільтрації [7] (зверху) та розрахований за двомасовою моделлю (знизу) (а); потік повітря u_g, обчислений для нейтрального (пунктирна лінія) та придихального (суцільна лінія) типу фонації (б); результати розрахунків потоку повітря $u_g(t)$ для трьох значень тиску на вході у голосову щілину $p_{s_1} = 0.02$, $p_{s_2} = 0.0012$, $p_{s_3} = 0.008$ (в); потік повітря u_g на виході з голосових зв'язок, обчислений з інерційними членами (суцільна лінія) та без них (пунктирна лінія) (г)

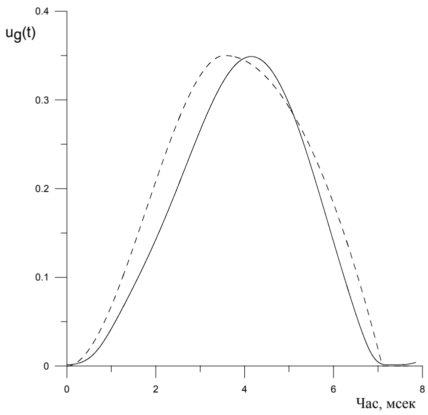


Рисунок 2 – Потік повітря u_g на виході з голосових зв'язок, обчислений за двомасовою моделлю (суцільна лінія), та R-моделлю (пунктирна лінія)

Для вивчення голосового джерела, окрім функції потоку повітря, використовується її похідна. Звичайний набір параметрів для кількісного опису коливань виводиться із залежності похідної потоку від часу і включає: амплітуду від'ємного піка, тривалість відкриття, тривалість закритої фази, проміжок часу між додатним і від'ємним піками, час повернення в початкове положення.

Проведений порівняльний аналіз отриманої чисельної похідної потоку повітря за двомасовою моделлю та похідної потоку повітря, отриманою за аналітичною LF-моделлю [8]:

$$\frac{du_g(t)}{dt} = \begin{cases} E_0 e^{\alpha t} \sin(\omega_g t), & 0 \le t < t_e, \\ -\dfrac{E_e}{\varepsilon t_a}\left[e^{-\varepsilon(t-t_e)} - e^{-\varepsilon(t_c-t_e)}\right], & t_e \le t < t_c, \\ 0, \quad t_c \le t < t_0, \end{cases}$$

де t_0 – період; t_e – відмітка мінімального значення похідної; t_a – визначається як точка, в якій тангенс перетинає вісь координат; t_c – момент, в який похідна під час фази закриття майже досягає нуля; E_e – абсолютне значення мінімуму похідної; α, ε, ω_g – деякі параметри. Результати порівняння моделей наведені на рис. 3.

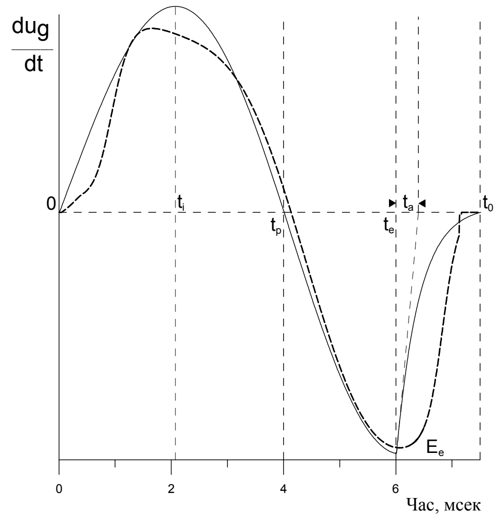


Рисунок 3 – Чисельна похідна від потоку повітря, обчисленого за двомасовою моделлю (пунктирна лінія), та похідна, обчислена за LF-моделлю (суцільна лінія)

Моделювання мовного тракту людини

Для моделювання розповсюдження акустичних хвиль у мовному тракті як у неоднорідній акустичній трубі, що починається між голосовими зв'язками та закінчується губами, використовується система рівнянь акустики в частинних похідних, яка записується у вигляді [9]:

$$-\frac{\partial p}{\partial x} = \frac{\rho}{S(x)}\frac{\partial u}{\partial t}, \qquad -\frac{\partial u}{\partial x} = \frac{S(x)}{\rho c^2}\frac{\partial p}{\partial t}, \tag{8}$$

де $0 \le x \le L,\ t > 0$, L – довжина мовного тракту; $p(x,t)$ – тиск у тракті в момент часу t; $u(x,t)$ – об’ємна швидкість потоку; ρ – густина повітря в тракті; c – швидкість звуку; $S(x)$ – функція площі поперечного перерізу.

Оскільки тракт має неоднорідний поперечний переріз, він розбивається на циліндричні секції однакової довжини з постійною площею перерізу (рис. 4).

Як крайова умова на вході в тракт вибирається потік $u_g(t)$, знайдений з (1) – (3). Виходячи з цього, маємо крайову умову: $u(0,t) = u_g(t)$. На протилежному кінці тракту задаємо умову $p(L,t) = 0$.

Різницева задача для апроксимації системи рівнянь (3) будується на рознесеній сітці. Для розв’язання використовується явний метод «чехарда». Цей алгоритм має високу обчислювальну ефективність і може використовуватися для моделювання у реальному часі [10]. Результати роботи алгоритму наведені на рис. 5.

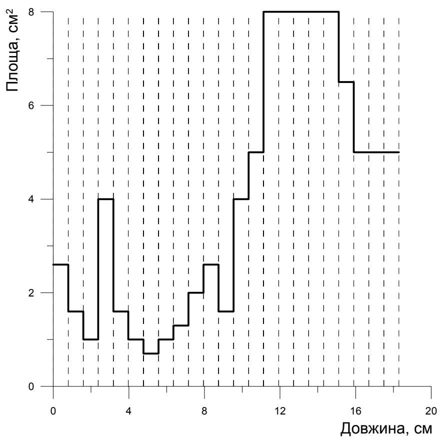


Рисунок 4 – Наближення мовного тракту циліндричними секціями однакової довжини

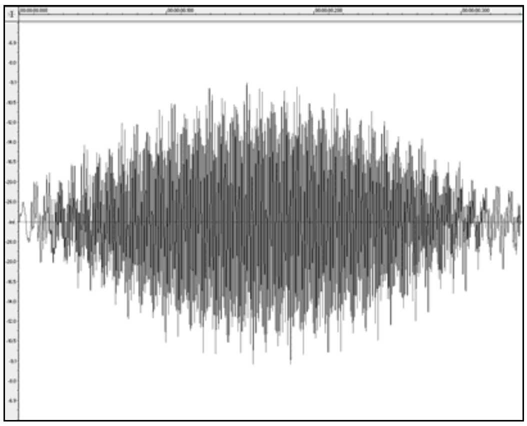


Рисунок 5 – Результати моделювання звуку [а] російської мови

Для моделювання розповсюдження акустичних хвиль також може використовуватися рівняння Вебстера [2]:

$$S(x)\frac{\partial^2 P}{\partial t^2} = c^2 \frac{\partial}{\partial x}\left(S(x)\frac{\partial P}{\partial x}\right), \tag{9}$$

де x – просторова координата уздовж середньої лінії тракту в середньо-сагітальній площині; t – момент часу; $p(x,t)$ – шуканий тиск у тракті; $S(x)$ – профіль площ поперечного перерізу вздовж тракту; c – швидкість звуку в тракті.

Як крайова умова на вході в тракт вибирається похідна від потоку повітря $P(0,t) = -\dfrac{\rho}{S(0)}\dfrac{du_g(t)}{dt}$. Для розв’язання задачі (9) використовується скінченнорізницевий метод. Для розв’язання системи різницевих рівнянь використовується ітераційний метод послідовної верхньої релаксації.

Розглянемо задачу відновлення форми мовного тракту за виміряними акустичними параметрами сигналу на базі акустичного рівняння Клейна – Гордона [11]. Для цього вводиться нова змінна $\varphi(x,t)$, яка визначається виразом:

$$\varphi(x,t) = P(x,t)S(x)^{1/2}. \tag{10}$$

Це дає можливість сформулювати акустичне рівняння у формі Клейна – Гордона:

$$\frac{\partial^2 \varphi(x,t)}{\partial t^2} = c^2 \frac{\partial^2 \varphi(x,t)}{\partial x^2} - c^2 U(x)\varphi(x,t), \; 0 < x < L \; 0 < t \le T. \qquad (11)$$

Рівняння (11) має форму хвильового, де функція $U(x)$ визначена в термінах площі поперечного перерізу мовного тракту як

$$U(x) = \frac{d^2 S(x)^{1/2} / dx^2}{S(x)^{1/2}}. \qquad (12)$$

Мовна обернена задача

Мовна обернена задача визначається як задача знаходження функції *S(x)* за виміряними параметрами мовного сигналу на виході з тракту. Математично ця задача розв'язується як задача пошуку мінімуму деякого функціонала за різного роду обмежень.

Нехай на виході з тракту вимірюється тиск $P(L,t)$, пов'язаний з розв'язком рівняння Клейна – Гордона співвідношенням (10). Позначимо $\Phi(t)$ функцію, яка вимірюється на виході з тракту.

Задача зводиться до мінімізації функціонала

$$J(U) = \int_0^T \left(\Phi(t) - \varphi_U(L,t)\right)^2 dt, \qquad (13)$$

де $\varphi_U(L,t)$ – розв'язок задачі (11) при заданій функції $U(x)$.

Для мінімізації функціонала (13) використовується градієнтний метод [12]. Приріст функціонала записується у вигляді:

$$\Delta J(U) = J(U+h) - J(U) = = \int_0^T 2(\Phi(t) - \varphi_U(L,t))\Delta\varphi \, dt + \int_0^T (\Delta\varphi)^2 dt,$$

де $\Delta\varphi = \varphi_{U+h}(x,t) - \varphi_U(x,t)$. Для визначення градієнта функціонала будується спряжена задача , яка записується у вигляді:

$$\frac{\partial^2 \Psi(x,t)}{\partial t^2} = c^2 \frac{\partial^2 \Psi(x,t)}{\partial x^2} - c^2 U(x)\Psi(x,t), \qquad 0 < x < L \; 0 < t \le T.$$

Градієнт функціонала визначається через розв'язок спряженої задачі за формулою

$$J' = -\varphi \, \Psi.$$

Після визначення $U(x)$ можна знайти $S(x)$ з (12).

У тестових розрахунках була використана наступна стратегія перевірки працездатності побудованого алгоритму та створеного програмного забезпечення. Розв'язувалась пряма задача та визначався сигнал на виході. Далі для розв'язування оптимізаційної задачі цей сигнал використовувався як виміряний. Для початкового наближення функції $U(x)$ використовувалась довільна функція, представлена графіком на рис. 6.

Точність розв'язку оцінюється за процедурою ресинтезу: синтезований за знайденим розв'язком сигнал має мало відрізнятися від вихідного сигналу, за параметрами якого вирішувалася обернена задача. Результати розрахунків наведені на рис. 7.

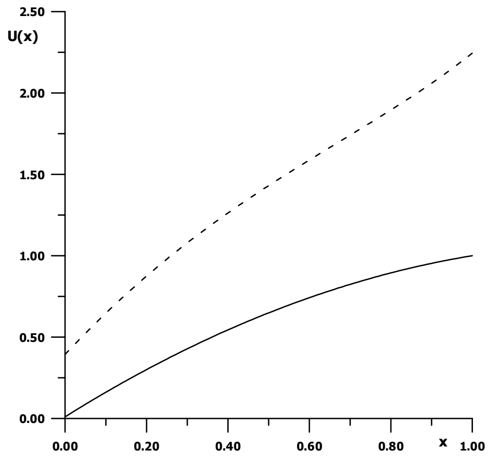


Рисунок 6 – Початкове наближення для функції $U(x)$ (суцільна лінія), уточнена функція $U(x)$ (пунктирна лінія)

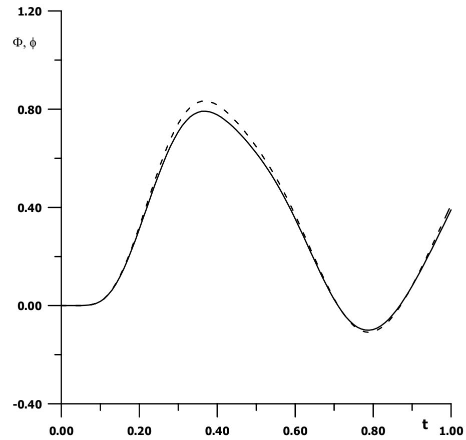


Рисунок 7 – Виміряний сигнал (суцільна лінія) $x = L$ та його наближення як розв'язок оберненої задачі (пунктирна лінія)

Програмно-алгоритмічний комплекс

Складність фізичного явища, що вивчається, призводить до необхідності урахування великої кількості параметрів. У свою чергу, математична модель є значною мірою чутливою до збурень цих параметрів. Для зручного введення та редагування вихідних параметрів моделі та відображення результатів моделювання був розроблений програмно-алгоритмічний комплекс. Екранна форма для введення параметрів моделі голосових зв'язок наведена на рис. 8.

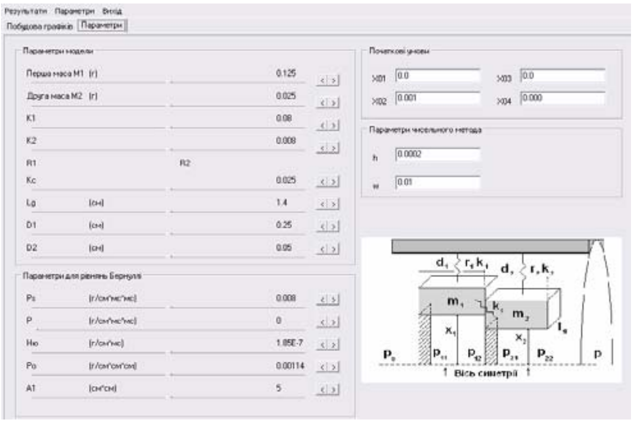


Рисунок 8 – Екранна форма для введення параметрів моделі голосових зв'язок

Необхідні для вивчення джерела параметри і результати моделювання відображаються у цифровому та графічному вигляді за допомогою екранних форм користувача.

Для вивчення голосового джерела, окрім функції потоку, використовується його похідна. Приклад відображення результатів моделювання потоку та його похідної наведений на рис. 9.

Для введення і редагування вихідних даних моделі мовного тракту, зокрема задання геометричної форми тракту та відображення результатів моделювання вихідного

мовного сигналу у цифровому та графічному вигляді, був розроблений графічний інтерфейс користувача, екранна форма якого наведена на рис. 10. За його допомогою користувач моделі може задавати бажану форму мовного тракту та вивчати її вплив на вихідний мовний сигнал (рис. 11).

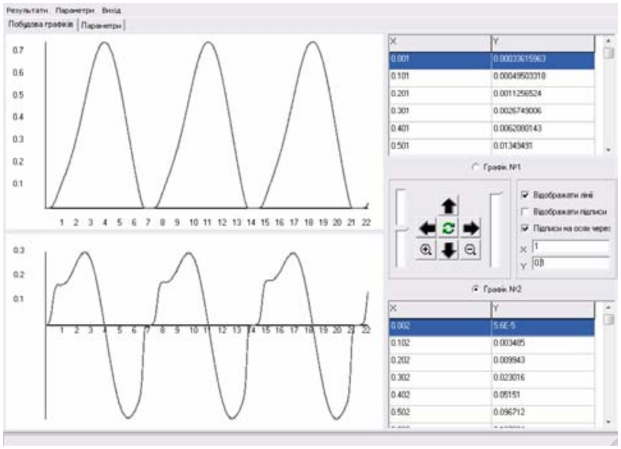


Рисунок 9 – Результати моделювання голосового джерела у графічному вигляді

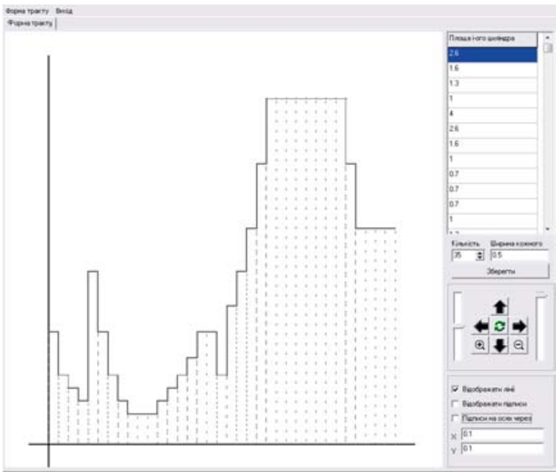


Рисунок 10 – Екранна форма для введення параметрів тракту

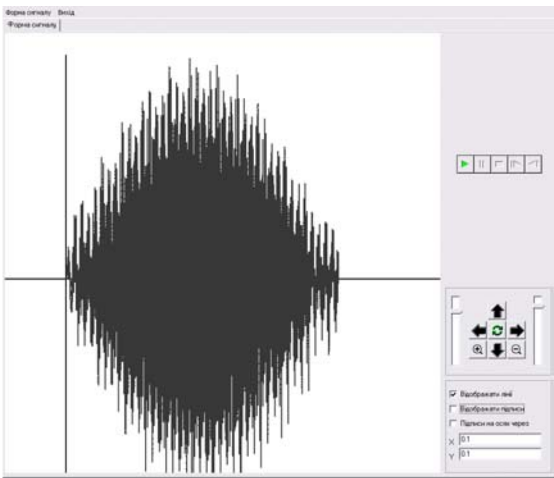


Рисунок 11 – Екранна форма для візуалізації вихідного сигналу

Висновки

У роботі запропонована і розроблена єдина система, яка дозволяє комп'ютерне відтворення звуків, що утворює людина на основі спільного використання фізичних моделей голосового джерела та мовного тракту. Також сформульовано і розв'язано обернену задачу відновлення параметрів мовного тракту. Подальші дослідження будуть направлені на удосконалення математичних методів та програмного забезпечення для моделювання голосових мовних сигналів.

Література

1. Леонов А.С. Обратная задача для управления артикуляцией / А.С. Леонов, В.Н. Сорокин // Доклады РАН. – 2000. – Т. 374, № 6. – С. 749-753.
2. Артикуляторный ресинтез фрикативных / Леонов А.С., Макаров И.С., Сорокин В.Н. [и др.] // Информационные процессы. – 2004. – Т. 4, № 2. – С. 141-159.
3. Ishizaka K. Synthesis of voiced sounds from a two-mass model of vocal cords / K. Ishizaka, J.L. Flanagan // Bell Syst. Tech. J. – 1972. – Vol. 51(6). – P. 1233-1268.
4. Арушанян О.Б. Численное решение обыкновенных дифференциальных уравнений на фортране / О.Б. Арушанян, С.Ф. Залеткин. – М. : МГУ, 1990. – 335 с.
5. Самарский А.А. Методы решения сеточных уравнений : [учеб. пособие для студентов вузов по спец. «Прикладная математика»] / А.А. Самарский, Е.С. Николаев. – М. : Наука, 1978. – 592 с.
6. Rosenberg A. Effect of glottal pulse shape on the quality of natural vowels / A. Rosenberg // Journal of the Acoustical Society of America. – 1971. – Vol. 49. – P. 583-590.
7. Rothenberg M. A new inverse-filtering technique for deriving the glottal airflow waveform during voicing / M. Rothenberg // J. Acoust. Sec. Am. – 1973. – Vol. 53, № 1. – P. 1632-1645.
8. Fant G. The lf-model revisited. Transformations and frequency domain analysis / G. Fant // STL-QPSR, Tech. Rep. – 1995. – Vol. 2-3. – P. 119-156.
9. Fundamentals of Acoustics / [Kinsler L.E., Frey A.E., Coppens A.B., Saunders J.V.]. – San Diego : Academic Press, 1982. – 496 p.
10. Van den Doel K. Real-time numerical solution of Webster's equation on a non-uniform grid / K. van den Doel, U. Ascher // IEEE Trans. Audio, Speech and Language Processing. – 2008. – Vol. 16. – P. 1163-1172.
11. Forbes B.J. The accoustical Klein-Gordon equation: The wave-mechanical step and barrier potential functions / B.J. Forbes, E.R. Pike, D.B. Sharp // J. Acooust. Soc. Am. – 2003. – № 114(3). – P. 1291-1302.
12. Васильев Ф.П. Методы решения экстремальных задач / Васильев Ф.П. – М. : Наука, 1981. – 400 с.

Ю.В. Крак, И.О. Стеля

Моделирование речевого аппарата человека в задачах компьютерного синтеза

В статье дальнейшее развитие получила методика решения проблемы синтеза речевой информации, что позволило решить задачу искусственного образования звуков голоса человека, основываясь на совместном использовании физических моделей голосового источника и речевого тракта. На базе разработанных алгоритмов создан программно-алгоритмический комплекс для исследования влияния различных параметров моделей в системе «голосовой источник – речевой тракт» на речеобразование. Для акустического уравнения Клейна – Гордона решена обратная задача восстановления параметров речевого тракта по измеренному сигналу на выходе.

Y.V. Krak, I.O. Stelia

Modeling Human Speech Apparatus in Problems of Computer Synthesis

The paper has further developed the methods of solving the issue of synthesizing the voice data, thereby solving the issue of artificial formation of sounds of a human voice based on the shared use of physical models of the voice source and vocal tract. On the basis of the developed algorithms there was created the program and algorithmic complex for studying of influence of parameters of the models in the language source – vocal tract system on speech production. For acoustical Klein – Gordon's equation the inverse problem of recovering the parameters of the vocal tract according to the measured output signal was solved.

Стаття надійшла до редакції 08.07.2010.